

Michigan Law Review

Volume 91 | Issue 6

1993

Intellectual History, Probability, and the Law of Evidence

Peter Tillers

Benjamin N. Cardozo Law School, Yeshiva University

Follow this and additional works at: <https://repository.law.umich.edu/mlr>



Part of the [Evidence Commons](#), and the [Legal History Commons](#)

Recommended Citation

Peter Tillers, *Intellectual History, Probability, and the Law of Evidence*, 91 MICH. L. REV. 1465 (1993).

Available at: <https://repository.law.umich.edu/mlr/vol91/iss6/25>

This Review is brought to you for free and open access by the Michigan Law Review at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Michigan Law Review by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

INTELLECTUAL HISTORY, PROBABILITY, AND THE LAW OF EVIDENCE

Peter Tillers*

"BEYOND REASONABLE DOUBT" AND "PROBABLE CAUSE": HISTORICAL PERSPECTIVES ON THE ANGLO-AMERICAN LAW OF EVIDENCE. By *Barbara J. Shapiro*. Berkeley: University of California Press. 1991. Pp. xv, 365. \$42.50.

INTRODUCTION

Revolutions are proclaimed more often than they occur. Nonetheless, we seem to be in the midst of a "probabilistic revolution."¹ This revolution goes beyond physics and Heisenberg's well-known Uncertainty Principle,² and it goes far beyond the relativistic nihilism of Critical Legal Studies. The probabilistic revolution extends to a wide variety of academic fields, including the social sciences,³ decision theory,⁴ biology,⁵ economics,⁶ logic,⁷ and philosophy of science.⁸ This

* Professor of Law, Director, International Seminar on Evidence in Litigation, Benjamin N. Cardozo School of Law, Yeshiva University. A.B. 1966, Yale; J.D. 1969, LL.M. 1972, Harvard. — Ed. I thank John Langbein, Richard Lempert, and David Schum for their comments. Errors and omissions are my own.

1. See generally THE PROBABILISTIC REVOLUTION (Lorenz Krüger et al. eds., 1987); STEPHEN M. STIGLER, THE HISTORY OF STATISTICS: THE MEASUREMENT OF UNCERTAINTY BEFORE 1900 (1986).

2. See, e.g., MICHAEL AUDI, THE INTERPRETATION OF QUANTUM MECHANICS 1-17 (1973); JOHN GRIBBIN, IN SEARCH OF SCHRÖDINGER'S CAT: QUANTUM PHYSICS AND REALITY 119-20, 155-57 (1984); David C. Cassidy, *Heisenberg, Uncertainty and the Quantum Revolution*, SCI. AM., May 1992, at 106.

3. See generally 2 THE PROBABILISTIC REVOLUTION, *supra* note 1, at 7-131 (discussing probability and statistics in psychology and sociology in the nineteenth and twentieth centuries).

4. See, e.g., RICHARD JEFFREY, THE LOGIC OF DECISION (2d ed. 1983); DETLOF VON WINTERFELDT & WARD EDWARDS, DECISION ANALYSIS AND BEHAVIORAL RESEARCH (1986); Ward Edwards, *Dynamic Decision Theory and Probabilistic Information Processing*, 4 HUM. FACTORS 59 (1962).

5. See generally 2 THE PROBABILISTIC REVOLUTION, *supra* note 1, at 227-369 (concerning evolutionary biology). The probabilistic and statistical character of genetics was recognized long ago. See STIGLER, *supra* note 1, at 263-361. The random character of genetic processes plays a central role in debates about the forensic use of DNA evidence. See, e.g., Richard Lempert, *Some Caveats Concerning DNA as Criminal Identification Evidence: With Thanks to the Reverend Bayes*, 13 CARDOZO L. REV. 303 (1991).

6. See generally 2 THE PROBABILISTIC REVOLUTION, *supra* note 1, at 133-97.

7. See, e.g., Ronald R. Yager, *New Paradigms for Reasoning with Uncertain Information*, 13 CARDOZO L. REV. 1005 (1991); L.A. Zadeh, *Fuzzy Sets*, 8 INFO. & CONTROL 338 (1965).

8. See, e.g., Richard C. Jeffrey, *Valuation and Acceptance of Scientific Hypotheses*, 23 PHIL. OF SCI. 237 (1956), reprinted in RICHARD C. JEFFREY, PROBABILITY AND THE ART OF JUDGMENT 14 (1992).

revolution is not confined to the academy. For example, it has influenced the art of statecraft,⁹ the conduct of war and business, the tactics of election campaigning,¹⁰ and the practice of medicine.¹¹

The Anglo-American law of evidence may have anticipated the probabilistic transformation of contemporary social and political life.¹² For quite some time — at least for 100 years and, if Professor Barbara Shapiro¹³ is right, for at least 200 years — the governing assumption of this body of law has been that all or practically all facts are uncertain and that proof of facts is always or almost always a matter of probabilities.¹⁴ Legal scholarship has also long emphasized the probabilistic nature of judicial proof.¹⁵ In recent years scholarly interest in the topic of uncertain forensic proof has intensified. As the use of statistical evidence in courtrooms has increased, legal scholars have devoted more and more attention to formal theories of uncertainty. This “new evidence scholarship”¹⁶ has generated a large body of literature on “trial by mathematics” and statistical evidence.¹⁷

9. The use of statistics by governments and bureaucrats was an early development. *See, e.g.*, IAN HACKING, *THE TAMING OF CHANCE* 16-54, 189-199 (1990). Problems of “social engineering” spurred the development of both probability theory and statistical theory. For example, the “law of large numbers” was in part the result of ruminations about the optimal number of judges or jurors. *See id.* at 81-104; LORRAINE DASTON, *CLASSICAL PROBABILITY IN THE ENLIGHTENMENT* 342-68 (1988). There has been an effort to persuade intelligence analysts to think in overtly probabilistic terms. *See* DAVID A. SCHUM, *EVIDENCE AND INFERENCE FOR THE INTELLIGENCE ANALYST* (1987). How successful this effort has been I cannot say.

10. The influence of polling on election campaigns requires no documentation.

11. Ward Edwards, *Summing Up: The Society of Bayesian Trial Lawyers*, 66 B.U. L. REV. 937, 937 (1986) (describing a doctor who uses a computer when making diagnoses).

12. Some students of the history of probability and statistics suggest that problems of forensic proof helped to inspire the development of formal probability theory by Enlightenment theorists. *See* DASTON, *supra* note 9, at xiv-xv (“Enlightenment jurisprudence . . . contributed to the notion of rationality that the probabilists hoped to mathematize . . .”); *id.* at xvi (legal influences were particularly influential during the “prehistory” of mathematical probability); *id.* at 14-15; *see also supra* note 9. However, the imagination of these Enlightenment theorists was generally stirred by Continental rather than Anglo-American problems of forensic proof.

13. Barbara J. Shapiro is Professor of Rhetoric at the University of California at Berkeley.

14. *See, e.g.*, *Brown v. Schock*, 77 Pa. 471, 479 (1875); *Stevenson v. Stewart*, 11 Pa. 307, 308-09 (1849) (“[T]he competency of a collateral fact to be used as the basis of legitimate argument, is not to be determined by the conclusiveness of the inferences it may afford in reference to the litigated fact. It is enough if these may tend, even in a slight degree, to elucidate the inquiry, or to assist, though remotely, to a determination probably founded in truth.”).

15. The idea that proof is a matter of probabilities is implicit in the version of the relevance rule espoused by James Bradley Thayer in the nineteenth century. *See* James B. Thayer, *Presumptions and the Law of Evidence*, 3 HARV. L. REV. 141, 144-45 (1889); *see also* 1 JOHN H. WIGMORE, *EVIDENCE IN TRIALS AT COMMON LAW* § 10 (Tillers rev., 1983); 1A JOHN H. WIGMORE, *EVIDENCE IN TRIALS AT COMMON LAW* §§ 38-41 (Tillers rev., 1983) (citing numerous cases and much secondary literature endorsing the relevance rule and the principle that proof of facts is frequently or always inconclusive).

16. Richard Lempert invented the label “new evidence scholarship.” *See* Richard Lempert, *The New Evidence Scholarship: Analyzing the Process of Proof*, 66 B.U. L. REV. 439 (1986).

17. *See, e.g.*, Conference, *Decision and Inference in Litigation*, 13 CARDOZO L. REV. 253 (1991); Symposium, *Probability and Inference in the Law of Evidence*, 66 B.U. L. REV. 377 (1986). *See* the survey of similar literature before 1983 in 1A WIGMORE, *supra* note 15, § 37.1, at 1011-13; *id.* § 37.6.

Despite the long pedigree of probability in the law of evidence and in evidence scholarship, there are sharp disagreements today about the basic nature of forensic proof. Some legal scholars are deeply skeptical about, if not necessarily opposed to, the new evidence scholarship.¹⁸ New evidence scholars themselves disagree about the nature of forensic evidence, inference, and proof.¹⁹ Although the result has been a lively and interesting debate, some of the participants in this debate have staked out positions that appear to be irreconcilable. Sometimes the protagonists seem not even to understand each other. In situations such as this, it is often useful to consider how scholarly debates and discussions came to have the character that they do; a look into the past can put problems in a new light. Fortunately, in recent years scholars have produced several enlightening studies of the historical origins of our present-day conceptions of uncertainty and probability.²⁰ I had hoped that Shapiro's study of the origins of the law of evidence would shed still more light on the phenomenon of uncertainty. I regret to say that *Beyond Reasonable Doubt* did not live up to my expectations.

I. HISTORICAL SCHOLARSHIP AND EVIDENCE SCHOLARSHIP

Shapiro devotes a large part of her new book to the history of the reasonable doubt and probable cause standards in the "Anglo-American"²¹ criminal process. Shapiro calls her investigation "a member of the species intellectual history of which the history of legal doctrine is a subspecies" (p. 249). Thus, Shapiro examines the "intellectual baggage" that underlies the "talismans" of reasonable doubt and probable

18. See, e.g., Paul Bergman & Al Moore, *Mistrial by Likelihood Ratio: Bayesian Analysis Meets the F-Word*, 13 CARDOZO L. REV. 589 (1991); Kenneth W. Graham, Jr., "There'll Always Be an England": *The Instrumental Ideology of Evidence*, 85 MICH. L. REV. 1204 (1987) (reviewing WILLIAM TWINING, *THEORIES OF EVIDENCE: BENTHAM AND WIGMORE* (1985)); Leonard R. Jaffee, *Prior Probability — A Black Hole in the Mathematician's View of the Sufficiency and Weight of Evidence*, 9 CARDOZO L. REV. 967 (1988); Charles Nesson, *The Evidence or the Event? On Judicial Proof and the Acceptability of Verdicts*, 98 HARV. L. REV. 1357 (1985); Laurence H. Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 HARV. L. REV. 1329 (1971).

19. The new evidence scholarship is not a monolithic intellectual movement. Peter Tillers, *Introduction to Symposium, Probability and Inference in the Law of Evidence*, 66 B.U. L. REV. 381, 382-83, 389-90 (1986); Peter Tillers, *Prejudice, Politics, and Proof*, 86 MICH. L. REV. 768, 771 (1988).

20. See, e.g., DASTON, *supra* note 9; IAN HACKING, *THE EMERGENCE OF PROBABILITY: A PHILOSOPHICAL STUDY OF THE EARLY IDEAS ABOUT PROBABILITY, INDUCTION AND STATISTICAL INFERENCE* (1975); HACKING, *supra* note 9; JOHN H. LANGBEIN, *TORTURE AND THE LAW OF PROOF* (1977); STIGLER, *supra* note 1; WILLIAM TWINING, *THEORIES OF EVIDENCE: BENTHAM & WIGMORE* (1985); John H. Langbein, *The Criminal Trial Before the Lawyers*, 45 U. CHI. L. REV. 263 (1978); see also DAVID SCHUM, *EVIDENTIAL FOUNDATIONS OF COMPLEX INFERENCE* ch. 2 (forthcoming 1994).

21. Shapiro's focus is generally on the criminal process before 1800. Shapiro correctly observes that the differences between English and American criminal practice were not as great before 1800 as they are now. See p. 251.

cause.²² Shapiro's discussion of the origins and evolution of reasonable doubt and probable cause, however, is part of a more general historical argument. Shapiro maintains that a new empiricist philosophy crystallized in Britain late in the seventeenth century and made its presence felt in fields such as theology, natural science, history, and philosophy. She argues that this new empiricism (as well as other intellectual traditions) influenced the development of the Anglo-American law of evidence and that by the nineteenth century empiricist terminology and ideas dominated legal discourse about evidence. *Beyond Reasonable Doubt* describes how empiricist language and notions crept into legal literature. To show how empiricist language and principles gradually infiltrated and took over the law of evidence itself, Shapiro traces changes both in the criminal process and in a variety of legal doctrines.

The subject matter of *Beyond Reasonable Doubt* is vast. Building on her earlier work,²³ Shapiro surveys intellectual currents not only in the law, but also in fields such as science, religion, rhetoric, and philosophy. The geographical range of Shapiro's book is also wide. She discusses developments in Scotland, America, and Continental Europe as well as England. Furthermore, *Beyond Reasonable Doubt* covers a great deal of time. Shapiro's book emphasizes developments from 1500 to 1800, a period of three centuries. Shapiro also discusses matters such as trial by jury in the thirteenth and fourteenth centuries²⁴ and the Romano-canon legal tradition in the Middle Ages.²⁵ Altogether, *Beyond Reasonable Doubt* discusses developments over the course of more than half a millennium of Western history.²⁶

The sweeping proportions of *Beyond Reasonable Doubt* may be jus-

22. P. xii. Shapiro gives at least two different explanations for the significance of the reasonable doubt and probable cause standards. First, Shapiro sometimes emphasizes the idea that legal formulations of these standards affected actual legal practice. See, e.g., p. xii ("[T]here is some reason to believe that the saying did have some impact on the doing."). Second, Shapiro notes the ideological or "legitimizing" function of formulations of rules of evidence. See p. 252 ("[This book] is about the construction of doctrines that would legitimate a fact-finding regime that was obviously full of human error and on which the lives and liberties of many citizens depended."); see also p. 249 ("The development of legal ideas . . . has played an important role in shaping the ways in which legal systems work, or at least in shaping contemporary and subsequent perceptions of how they work." (emphasis added)).

23. BARBARA J. SHAPIRO, PROBABILITY AND CERTAINTY IN SEVENTEENTH-CENTURY ENGLAND: A STUDY OF THE RELATIONSHIPS BETWEEN NATURAL SCIENCE, RELIGION, HISTORY, LAW, AND LITERATURE (1983); Barbara J. Shapiro, *Law and Science in Seventeenth-Century England*, 21 STAN. L. REV. 727 (1969); Barbara J. Shapiro, "To a Moral Certainty": *Theories of Knowledge and Anglo-American Juries 1600-1850*, 38 HASTINGS L.J. 153 (1986).

24. See, e.g., pp. 3-4.

25. See, e.g., pp. 115-20.

26. Shapiro acknowledges that she takes a "long view." P. 250. Although she admits that "long-term approaches cannot provide the same detailed attention to context" that short-term approaches can, she argues that a long view is "essential for tracing the central intellectual dynamic of the history of the law of evidence" because "that dynamic is one in which ideas or doctrines created in one era and context are adapted and recycled in another." P. 250.

tified. Shapiro's historiographical tastes run toward intellectual history. It is notoriously difficult to pinpoint the exact place or time of important changes in the "thinking" of a society²⁷ or culture; intellectual traditions often originate in distant times and places and broad intellectual currents ordinarily change slowly. Shapiro's preference for intellectual history, however, does not fully explain the scope of her investigation. Shapiro examines developments after 1800 as well as developments before 1500. Of course, no particular year can serve as a definitive marker of the end of an intellectual tradition; like old soldiers, intellectual currents fade away slowly. However, Shapiro discusses developments that occurred long after 1800. For example, she considers the U.S. Supreme Court's treatment of probable cause in the 1980s (p. 147). Because Shapiro peers into the present as well as into the past, it is fair to infer that Shapiro wants to throw light on contemporary problems in the law of evidence as well as on the historical origins of the law of evidence.

The history of the law of evidence is an unjustly neglected field of scholarship. Shapiro's earlier work has already done much to fill this void.²⁸ *Beyond Reasonable Doubt* appears to be another important addition to scholarship in the history of the law of evidence. Shapiro's discussion of the intellectual background of the modern law of evidence contains a wealth of interesting detail, much of which is new to me. For example, I first learned from Shapiro about the possible influence of both the rhetorical tradition and English religious casuistry on the early English law of evidence.²⁹ Moreover, Shapiro's explanation of the medieval view of presumptions (pp. 213-41) is useful for amateur legal historians like myself who occasionally forage in older English reports for interesting discussions of evidence. However, since I am only an amateur in matters historical, I am not in a position to make a definitive judgment about *Beyond Reasonable Doubt* as a work of historical scholarship. Hence, in this essay I focus on the significance of *Beyond Reasonable Doubt* for contemporary evidence scholarship; I say little about its significance for historical scholarship.

My attitude of agnosticism about the extent of Shapiro's contribution to historical scholarship is not disingenuous; although I believe that *Beyond Reasonable Doubt* adds little to our understanding of contemporary problems of evidence and proof in the law or to our understanding of the nature of evidence and proof in general, I believe that my critique of the theoretical dimensions of Shapiro's book does not necessarily call into question Shapiro's contributions to legal and intellectual history. Nonetheless, my critique of *Beyond Reasonable Doubt*

27. Shapiro's concern is not so much with ideas in society at large, but with the opinions and ideas of "legal professionals and relatively educated persons." P. 252.

28. See *supra* note 23.

29. Pp. 13-18; see also SHAPIRO, *supra* note 23, at 78-86.

may speak to the *nature* of Shapiro's contributions to historical scholarship.

Although Shapiro's main interest is intellectual history and *Beyond Reasonable Doubt* is generally "short . . . on institutions" (p. 251), Shapiro does dabble in the history of legal institutions; she asserts that her general findings lie "at the boundary between intellectual and institutional history" (p. xv). As I explain in the remaining parts of this essay, Shapiro's theoretical arguments in *Beyond Reasonable Doubt* are unsatisfactory because they rest on theoretically unsophisticated notions about the nature of evidence, inference, probability, and proof. In one respect, the theoretical shortcomings of *Beyond Reasonable Doubt* affect Shapiro's historical arguments as well as her theoretical claims; they throw into question Shapiro's analysis of the causal relationship between institutional change and changes in the law of proof. Shapiro's discussion of the historical connection between the Anglo-American law of evidence and the institution of the trial jury is a case in point.

Shapiro explains the emergence of rules of evidence in criminal trials in England not by reference to intellectual currents and traditions, but by reference to changes in the way that the jury was expected to work. The modern trial jury is a passive body. It is not permitted to conduct its own investigation. It is not even permitted to make use of evidence that individual jurors chanced to acquire before they became jurors; generally speaking, the jury is permitted to consider only the evidence submitted by the parties at the trial. According to Shapiro and other historians, the medieval English trial jury was a "self-informing" body (p. 4). It was permitted and expected to conduct its own investigation, and it was permitted and expected to consider evidence and information within the personal knowledge of members of the jury. The idea that jurors may consider only the evidence that the parties present at trial was a later development, one that is completely antithetical to the medieval conception of the role of the jury. The medieval jury was expected to be the repository of the evidence and the facts; if the standard historical account holds water, no one would have imagined that the parties would have to inform the jury of the evidence and the facts.

Shapiro explains the emergence of rules of evidence in the English criminal process by reference to this transformation of the trial jury from a self-informing body to a passive receptacle for evidence submitted by others. Shapiro claims that rules of evidence became necessary when the jury ceased to be a self-informing body because at that point it became necessary for the jury to rely primarily on the testimony of witnesses rather than on the jury's own observations and investigation. She asserts that when the jury was a self-informing body "there had been no need to construct a rationale for the truth-finding capacities of

juries" (p. 11). Elsewhere she states: "A turning point occurred in 1563, when legislation compelled the attendance of witnesses and made perjury a crime. As witnesses became more important, juries increasingly required standards for the evaluation of testimony" (p. 6). Shapiro's reasoning implies that truth-seeking considerations explain the historical connection between the emergence of rules of evidence and the change in the role of the jury in the criminal process; she implies that the need for reliable assessment of testimonial evidence made rules of evidence necessary.

Shapiro's explanation of the birth of rules of evidence in medieval England is suspect. Shapiro acknowledges that even when the jury was a "self-informing" body it had to rely on sources of information other than the direct personal knowledge of its members.³⁰ If unsworn or informal witnesses as well as sworn or official witnesses can make mistakes and lie, the epistemic need for rules of evidence may be as great (or as little) when the jury is a self-informing body as when it is not. Moreover, even if one makes the unrealistic assumption that medieval juries acquired evidence solely by direct observation of non-testimonial evidence, jurors themselves are subject to "testimonial" infirmities such as weak eyesight and weak memory. Hence, if the weakness of testimonial evidence alone creates a need for rules of evidence, this need exists even if the jury obtains all of its information from nontestimonial sources.

Although it is possible that truth-seeking considerations explain why rules of evidence are more important when the jury is a relatively passive receptor of evidence than when it is an active, self-informing, and investigative body, Shapiro had to say more than she did to explain why that is so.³¹ A more satisfactory account of the origins and persistence of rules of evidence would consider not just the role of the trial jury, but also matters such as the procedural rules governing the trial and pretrial process, the role of actors such as the trial judge, the interests of participants in litigation, the importance attached to the accuracy of the proof process, and the ordinary, or natural, inferential methods of fact finders.

Professor John Langbein's work illustrates the kind of research and analysis that is required to explain the origins of the rules of evi-

30. P. 4. Shapiro observes that recent studies have challenged the assumption that the jury was a fully "self-informing" body. P. 4. Even if such studies did not exist, one could safely assume that the self-informing jury did not get all of its information from its own first-hand observations; even if self-informing juries conducted their own factual investigation, they must have gotten some of their information from nonjurors.

31. Although it may be unwise to require investigative bodies to adhere to formal rules of evidence, see 1 WIGMORE, *supra* note 15, § 4, at 53-54, the folly of such a requirement is not self-evident. For example, the probable cause standard applies to grand juries. Moreover, some states today ostensibly require that grand juries rely only on "legal evidence" when making a probable cause determination. *Id.* at 56 n.34. *But cf. id.* at 59 n.35 (The Supreme Court recognized that federal grand juries are not bound to follow ordinary rules of evidence.).

dence. For example, Langbein's work speaks to the influence of the interests of participants in litigation on the law of proof; in one paper he takes the position that the existence of rules of evidence has more to do with the participation of lawyers in the adjudicative process than with the institution of trial by jury.³² Langbein's work also illustrates why it matters how much importance a society or a legal system attaches to the accurate determination of facts in adjudication. Although Langbein emphasizes the corrupting influence of the interests and ideologies of groups such as lawyers, he also assumes that bankrupt proof schemes tend to break down precisely because they are bankrupt: an epistemologically impoverished proof regimen begins to collapse when its poverty becomes apparent.³³ This thesis presupposes that society, or a relevant segment of it, sees accurate fact finding as a central function of its system of adjudication. Finally, although Langbein does not directly address the question of how ordinary inferential processes work, his arguments suggest that this question is also important. Langbein's belief that corrupt proof systems tend to break down when their corruption becomes apparent assumes that there is a tension between inadequate, legally prescribed methods of fact finding and sensible, natural methods of finding facts, a tension that eventually becomes too great for a society to bear. Whether or not Langbein is right to assume that ordinary, or natural, inference works well, this question is relevant to hypotheses about the historical origins of the rules of evidence.³⁴

II. PROBABILITY AND PROBABILITY THEORY

Shapiro's book ends with a short appendix on mathematical theories of probability. In this appendix she observes: "The frequent reference to the concept of probability [in this book] no doubt reminds the reader of mathematical notions of probability and of contemporary discussion of the role of mathematical probability and the law" (p. 253). Shapiro is correct. Many of the arguments and theories that she surveys and many of Shapiro's own discussions are reminiscent of mathematical theories of probability. This is not surprising. The emergence of a new concept of probability is one of the central themes of Shapiro's work.³⁵ Nevertheless, Shapiro's short appendix is her

32. Langbein, *supra* note 20, at 306.

33. *Id.* at 300-16; John H. Langbein, *Torture and Plea Bargaining*, 46 U. CHI. L. REV. 3, 12-22 (1978).

34. In recent years there have been important empirical studies of how natural, or ordinary, factual inference works. See, e.g., REID HASTIE ET AL., *INSIDE THE JURY* 151-58 (1983); Margaret Bull Kovera et al., *Jurors' Perception of Eyewitness and Hearsay Evidence*, 76 MINN. L. REV. 703 (1992); Nancy Pennington & Reid Hastie, *A Cognitive Theory of Juror Decision Making: The Story Model*, 13 CARDOZO L. REV. 519 (1991) (summarizing results of prior research). Perhaps historians of the law of evidence should examine some of this literature.

35. See, e.g., p. xiii (*Beyond Reasonable Doubt* is about "how changing conceptions of probability and certainty helped to shape legal formulations.").

only significant discussion of formal probability theory.

Shapiro's decision to give short shrift to formal probability theory may have been unwise. Modern theories of evidence, inference, and proof — including mathematical theories — may be relevant to the historical study of conceptions of evidence and proof. Moreover, interest in mathematical approaches to problems of forensic proof is not new. Shapiro herself mentions various English and American legal scholars and intellectuals who discussed the use of numbers and mathematics in argument about evidence (pp. 253-55). Although Shapiro does not explain why she ignores mathematical theories even when she examines the work of these theorists,³⁶ her brief comments in the appendix suggest that she felt free to do so because British and American legal scholars generally rejected mathematical and numerical methods.³⁷ This, however, is not a sufficient reason for ignoring mathematical theories of probability. If the people whose thinking we wish to decipher believed that mathematical methods merit attention, the methods they considered (and rejected) should also interest us. The precise grounds for their rejection of mathematical analysis can tell us much about their views about evidence and proof in adjudication. For example, it would be useful to know whether they objected to the use of numbers because they believed that (a) mathematical theories of probability are "wrong," (b) decision makers such as jurors do not know how to compute probabilities by using numbers, (c) the use of numbers requires excessively precise probability estimates, or (d) moral, social, or institutional considerations make it important to allow decision makers such as jurors to make judgments about facts in an ordinary way. Or did they reject mathematical analysis simply because (e) they did not understand mathematical probability or because (f) they sensed that the mathematical theories of probability of their day were not rich enough to capture the true complexity of problems of evidence in the law?

To determine why British and American intellectual figures in the seventeenth, eighteenth, and nineteenth centuries rejected mathematical analysis of evidence, it may be necessary to study the mathematical

36. Shapiro acknowledges the work of historians of probability such as Lorraine Daston and Ian Hacking. See p. 329 n.1. Daston has much to say about the historical relationship between legal conceptions of proof and mathematical theories of probability. See DASTON, *supra* note 9, at xiv-xvi, 14-15, 342-68; see also HACKING, *supra* note 9.

37. Shapiro's appendix on probability theory apparently does not exaggerate the general aversion of American and English historical scholars to mathematical and numerical analysis of evidence. However, William Best was a notable exception. See 1 WILLIAM BEST, *THE PRINCIPLES OF THE LAW OF EVIDENCE* 69-75 (James A. Morgan ed., James Cockcroft & Co., 1st Am. ed. 1875) (1849). Moreover, some English intellectuals were prominent players in the Enlightenment movement that advocated the use of the probability calculus in moral contexts law. See DASTON, *supra* note 9, at xvi, 296-369.

Legal authors generally heaped scorn on Bentham's "thermometer of persuasion." TWINING, *supra* note 20, at 59-60.

theories of probability that were in vogue at the time. Lorraine Daston has shown that between 1650 and 1840 theorists in a variety of fields attempted to apply the "calculus of chances" to "practical" and "moral" problems, including problems of forensic proof.³⁸ The British and American thinkers who scoffed at mathematical analysis of evidence were probably aware of these attempts. Thus, they may have had their mathematically minded contemporaries and predecessors in mind when they argued against the use of mathematical methods in law and in other "moral sciences."³⁹

Formal probability theory may be relevant to the historical study of Anglo-American ideas about evidence and proof for other reasons. A distinctive characteristic of modern conceptions of proof is the assumption that factual proof is a matter of probabilities and degrees. Shapiro maintains, at least for the most part,⁴⁰ that almost all British and American intellectuals who developed and propagated the new empiricism embraced this pivotal assumption. If Shapiro is right, these intellectuals may have embraced a basic premise of a twentieth-century theory of uncertainty, the standard probability calculus.

In the standard calculus of probabilities, the probability of a hypothesis can range from "impossible" to "certain." The probability of a hypothesis is represented by real numbers. By convention "0" represents impossible; "1," certain; and numbers between "0" and "1," degrees of probability between impossible and certain, with larger numbers representing greater degrees of probability. In this system probability can have an infinite number of values because the set of real numbers in the interval from "0" to "1" is uncountably infinite. By proclaiming that factual proof involves probabilities and degrees of probabilities, Anglo-American thinkers may have effectively accepted the tenet of the standard modern theory of probability that uncertainty is properly gauged or expressed by a scale consisting of a continuum of values.

If a person grants that probabilities are part of a continuum, it does not follow that he thinks about uncertainty the same way that a proponent of the standard probability calculus does. The standard calculus works as it does only if certain other premises are granted.⁴¹ Nonethe-

38. DASTON, *supra* note 9, at xvi, 40-47, 296-369. Most of the probabilists who advocated the use of calculus in the law were not English. However, at least one English religious figure used probability calculus in an attempt to justify the credibility of miracles and religious tradition. *Id.* at 312-15 (John Craig's theory of the probability of testimony).

39. John Stuart Mill, who rejected numerical analysis of forensic evidence, explicitly referred to proposals for using the doctrine of chances to assess the credibility of witnesses. *See* DASTON, *supra* note 9, at 297; *see also* p. 254. Shapiro reports that William Wills was also familiar with such literature. P. 254.

40. *See infra* text accompanying note 47.

41. For example, the probability calculus requires the assumption that the probability of a hypothesis and the probability of its negation are complementary or, more generally speaking, the assumption that the probabilities of disjoint and exhaustive hypotheses sum to one. *See* IAN

less, there are several reasons why the standard theory of probability may be a useful heuristic device for historical investigation of conceptions of evidence and proof. The first and most obvious reason is the possibility that at least some British and American empiricists effectively embraced the other premises that make the standard probability calculus run. Second, even if analysis shows that some historical figures did not accept all of the central premises of the standard modern theory of probability, a comparison of the principles favored by such historical figures and those embraced by conventional modern probabilists could throw historical conceptions of evidence and proof into sharper relief and thus make them more transparent and more intelligible to a modern observer. Finally, careful inspection of any differences between the assumptions of modern probabilists and those of historical figures might show that some historical conceptions of evidence and proof resemble one of the several nonstandard formal theories of uncertainty that exist today; that is, investigation might show that some British or American intellectual traditions departed from the assumptions underlying the standard calculus in much the same way that a particular nonstandard formal theory of uncertainty does. This in itself would be an important discovery because it would open new lines of investigation.⁴²

My claim that the standard theory of probability may have heuristic value in historical investigation does not presuppose that the standard probability calculus is the only valid paradigm of inferential argument. The standard calculus is only one of several formal theories of uncertainty.⁴³ However, Shapiro fails to make use of any of these formal nonstandard theories. For example, she has little to say about L. Jonathan Cohen's theory of eliminative induction.⁴⁴ Cohen maintains that his theory, which ascribes ordinal rather than cardinal properties to uncertain judgments about facts, has its wellsprings in English empiricism.⁴⁵ Since one of Shapiro's main interests is the influence of English empiricism on the law of evidence, her failure to

HACKING, *THE LOGIC OF STATISTICAL INFERENCE* 19 (1965). In a generalized version of the standard calculus, the principle that the probabilities of disjoint and exhaustive hypotheses must sum to one is viewed as a special case of the principle that where $A, B, C, \dots N$ are disjoint events, the probability of either A, B, C , or $\dots N$ is the sum of the individual probabilities of $A, B, C, \dots N$. In short: $P(A \text{ or } B \text{ or } C \text{ or } \dots N) = P(A) + P(B) + P(C) + \dots P(N)$.

42. Lorraine Daston distinguishes between conceptual probability and mathematical probability. DASTON, *supra* note 9, at 13. The distinction is necessary because, as Daston writes, some conceptions of probability are less amenable to mathematical interpretation than others. However, the distinction does not entirely justify jettisoning the use of mathematical probability in the study of the history of conceptions of evidence and proof. Some conceptions of probability are amenable to mathematical treatment.

43. Tillers, *supra* note 19, at 382-83, 389-90.

44. L. JONATHAN COHEN, *THE PROBABLE AND THE PROVABLE* (1977).

45. *Id.* § 12, at 42; *id.* § 46, at 144-51 ("Mill's canons" can be subsumed under Cohen's "method of relevant variables."). Cohen calls his system of probability "Baconian" in honor of Francis Bacon. *Id.* § 12, at 42-43.

make use of Cohen's "Baconian" theory may be even more unfortunate than her failure to make use of the standard probability calculus.⁴⁶

Although my criticism of Shapiro's treatment of formal theories of uncertainty stands on its own, it also highlights a more general weakness of Shapiro's investigative strategy. The overarching flaw in Shapiro's methodology is her failure to consider the epistemic needs and interests of a significant part of her intended audience — students of evidence. Parts III and IV of this essay spell out this criticism. Part V suggests how Shapiro might have structured her historical investigation to make it more relevant to contemporary debates about the nature of forensic evidence and proof.

III. PROBABILITY AND PRACTICAL CERTAINTY

Shapiro's treatment of the reasonable doubt standard illustrates how failure to attend to theoretical principles and distinctions can damage historical investigation of conceptions of evidence and proof. Shapiro reports that as the influence of British empiricism grew, members of the Anglo-American intelligentsia increasingly insisted that factual proof has a probabilistic character and always falls short of the absolute certainty that strict logical demonstration provides.⁴⁷ Despite this, Shapiro asserts, many of these same people insisted that proof of facts to a "practical certainty" is possible.

The idea that there is a practical certainty short of absolute certainty may strike some observers as odd. The idea seems incoherent at first blush because it seems tantamount to the notion of "certain uncertain proof," which is a contradiction in terms. Shapiro is not oblivious to the problematic character of the notion of practical certainty. She describes various attempts to explain and vindicate the notion, and she pays particular attention to an explanation advanced by John Locke and some other empiricists. However, as we shall see, in the eyes of many modern observers, these explanations may be as incoherent as the concept of practical certainty itself.⁴⁸

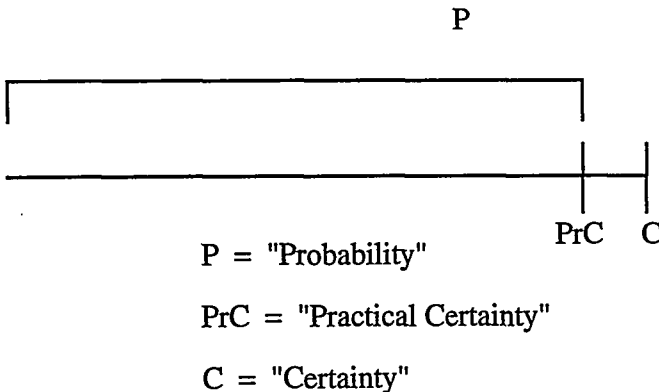
46. See *infra* notes 52-56 and accompanying text for further discussion of Cohen's theory and its uses.

47. See, e.g., p. 33. Shapiro sometimes hedges her claim that the new British thinkers believed that proof to a practical certainty falls short of absolute certainty. For example, she does so in her discussion of Thomas Reid and the Scottish "Common Sense" school of philosophy. See, e.g., p. 28. Shapiro does not view Reid's position as an aberration. She asserts that Reid and the Common Sense school had great influence on legal treatises on evidence. P. 223.

48. It is not clear whether Shapiro personally endorses either the concept of practical certainty or the explanations of practical certainty that theorists such as John Locke provided. However, the theoretical significance of *Beyond Reasonable Doubt* may not depend on whether Shapiro is an advocate as well as an observer. Consider Shapiro's stance toward Locke's theory. See *infra* text accompanying notes 50-51. If Locke's interpretation of reasonable doubt and practical certainty is incoherent and Shapiro accepts Locke's interpretation, Shapiro's own position is incoherent. If Locke's interpretation is incoherent but Shapiro does not endorse it, Shapiro's

According to Shapiro, some English theorists — most notably, John Locke — believed that proof to a “practical (or moral) certainty” is tantamount to proof to “the highest degree of probability” (p. 8). These theorists apparently believed that the notion of the highest probability provides a valid basis for a distinction both between practical certainty and absolute certainty and between practical certainty and lesser degrees of probability. For example, many of these empiricists subscribed to the view that matters such as “opinion” and “suspicion” belong to the realm of “conjecture” and “mere” probability, and they generally admitted that practical certainty falls short of absolute certainty. However, these empiricists also insisted that proof to a practical certainty is something more than a merely probable proof: proof to a practical certainty establishes a fact to the highest degree of probability whereas a merely probable proof does not. Thus, if Shapiro’s account is accurate, John Locke and the other British thinkers she discusses envisioned proof in the manner shown by Figure 1.

FIGURE 1

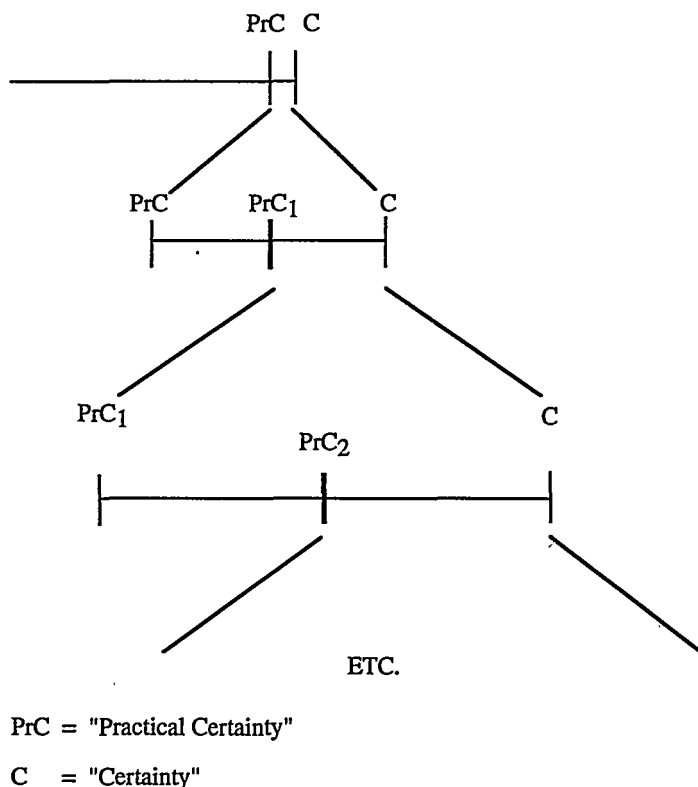


The idea of a highest possible probability short of absolute certainty is not without its difficulties. For example, a theorist who regards the standard probability calculus as the only coherent interpretation of uncertainty must insist that the only coherent notion of certainty is complete or absolute certainty. She must therefore reject the concept of a “highest degree of probability” as incoherent. Consider the interval $\{PrC, C\}$ in Figure 1. Suppose that this interval represents the difference between “practical certainty” and “certainty.” This interval, no matter how small, is a finite interval. If every probability is a point on a continuum, the interval PrC to C can always be made smaller; any finite interval is infinitely divisible.

historical analysis may be bereft of theoretical significance. Has Shapiro been caught by this kind of dilemma? My answer is a qualified “yes.”

Therefore, PrC cannot possibly be the highest value. Thus, no matter how close PrC is to C , it can be made closer. For example, as Figure 2 shows, any such interval can always be cut in half. An uncountably

FIGURE 2



large number of points can lie in any interval between C and any point short of C , no matter how small that interval is. Hence, if the conventional theory of probability describes how probability behaves, a proof that fails to make a hypothesis certain (or, alternatively, impossible) cannot possibly prove a fact to the "highest degree of probability."⁴⁹

49. Cf. Jeffrey, *supra* note 8, at 1-2. Jeffrey states:

The [Bayesian] framework . . . replaces the two Cartesian options, affirmation and denial, by a continuum of judgmental probabilities in the interval from 0 to 1, endpoints included, or — what comes to the same thing — a continuum of judgmental odds in the interval from 0 to ∞ , endpoints included. Zero and 1 are probabilities no less than 1/2 and 99/100 are. Probability 1 corresponds to infinite odds, 1:0. That's a reason for thinking in terms of odds: to remember how momentous it may be to assign probability 1 to a hypothesis. It means you'd stake your all on its truth, if it's the sort of hypothesis you can stake things on.

Id. However, Jeffrey errs in saying that infinity is an endpoint. Infinity has no end.

Practitioners of formal probability theory are not the only people who might make the kind of objection I have described. A perfectly "ordinary" student of evidence might make a similar objection in the following way:

I do not understand how there can be a "highest probability." Suppose that the available evidence seems to make a fact very probable and thus practically certain. Regardless of how probable any such fact may seem, it is always possible to imagine an additional piece of evidence that will make a possible fact seem more probable. Hence, there is no highest possible or imaginable probability of a fact in issue.

It is unclear whether Shapiro believes that the idea of practical certainty is coherent.⁵⁰ Hence, if Shapiro were faced with the sort of criticism that I have just described, she might respond that her job is to report past beliefs about proof, not to defend them. However, this would be an inadequate response.⁵¹ Shapiro tacitly promised not just to show us how people once thought about proof, but also to shed light on contemporary problems of proof. In any event, Shapiro's history of ideas about reasonable doubt and practical certainty may give our intellectual ancestors less credit than they deserve.

Consider an example. Shapiro maintains that the founders of the modern theory of evidence believed that proof of facts involves degrees of probability. This amounts to the claim that those founders believed that proof should be graded on a cardinal scale. This means, for example, that they believed that one proof might be twice as strong as another. However, Shapiro does not consider the possibility that the historical figures she discusses believed that it is sometimes appropriate to grade proof with an ordinal scale. This would mean that they believed that sometimes it is permissible to say only that one proof is stronger or weaker than another proof, but not how much stronger or weaker.

L. Jonathan Cohen has developed a sophisticated and well-known ordinal theory of probability and proof.⁵² Like some of the historical figures that Shapiro discusses, Cohen maintains that something akin to practical certainty about facts is possible.⁵³ My criticism of the "highest probability" interpretation of the reasonable doubt standard and

50. The passage quoted *infra* in the text accompanying note 57 suggests that Shapiro thinks that the idea of practical certainty is both coherent and sensible.

51. See *supra* note 48.

52. COHEN, *supra* note 44.

53. *Id.* §§ 27-28, at 82-86; § 68, at 247-252; *id.* at 265 ("Proof beyond reasonable doubt is proof at the level of inductive certainty."); see also 1 SCHUM, *supra* note 9, at 106-11, 397-401 (explaining Cohen's theory of "weight" and "provability"); see also L. Jonathan Cohen, *The Role of Evidential Weight in Criminal Proof*, 66 B.U. L. REV. 635 (1986). There is an important difference between the concept of "certainty" in standard probability theory and the concept of "inductive certainty" in Cohen's Baconian probability system. In Cohen's system a hypothesis that has been proven is subject to disproof by additional evidence. In the standard probability calculus a proposition that is certain cannot be refuted or weakened by evidence.

practical certainty does not apply to Cohen's ordinal interpretation of the reasonable doubt standard.⁵⁴ It is unfortunate that Shapiro does not make use of Cohen's theory.⁵⁵ Cohen maintains that his theory lies squarely in the intellectual tradition epitomized by British empiricists such as Bacon, Hume, and J.S. Mill.⁵⁶ If Cohen is right, British empiricists make sense only if their words are read through ordinal lenses.

The difficulties that Shapiro encounters because of her failure to use nonstandard theories of probability and proof are not limited to her discussion of the reasonable doubt standard. As the next part of this essay demonstrates, Shapiro's analysis of the probable cause standard encounters a variety of problems. As in other contexts, Shapiro assumes that factual proof concerns only degrees or levels of probabilities. However, her analysis of the vagaries of the probable cause standard also rests on the thesis that there are fundamental differences among various types of probabilistic proof. This thesis may be incompatible with the assumption that proof involves only probabilities.

IV. PROBABILITY AND PROBABLE CAUSE

Shapiro argues that the probable cause standard, in contrast to the reasonable doubt standard, has been unstable for centuries. She explains this in the following way:

The realm of probability is a clearly intermediate one between absolute opinion and mere opinion or rumor, but it contains no clear steps, degrees, or levels of probability in the scale of just above opinion to just below certainty, and it has no agreed upon measure of quantities of probability. . . . Purely epistemological discussion of grand jury standards [standards which require more than mere opinion but something less than almost complete certainty] is inherently unstable. It constantly moves up the scale toward "satisfied conscience," "moral certainty," and "beyond reasonable doubt," or down the scale toward "suspicion" or "opinion" because there is no fixed intermediate point at which to rest.
[p. 44]

Shapiro's argument assumes several distinct levels or species of proof. The strongest type of proof is absolute or demonstrative proof, which lies at one end of the spectrum. Short of absolute proof, but close to it, is practical certainty. Further removed is everything else,

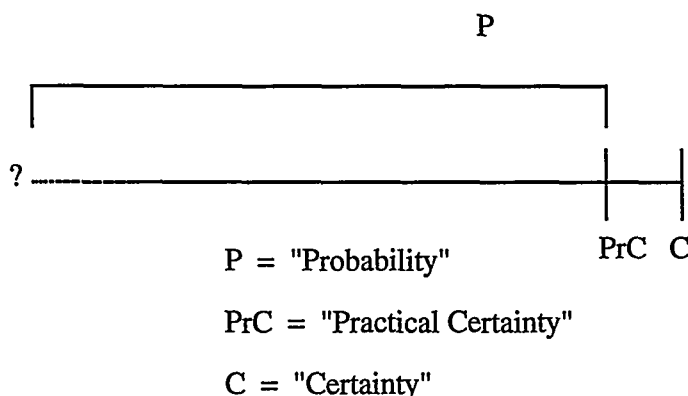
54. An ordinal scale merely ranks probabilities. Unlike a cardinal scale, an ordinal measure of probabilities cannot express the extent or magnitude of the difference between any two probabilities. Hence, an ordinal scale does not rule out the idea of a highest probability short of certainty.

55. Although there is a citation to Cohen in the bibliography for *Beyond Reasonable Doubt*, Shapiro does not discuss Cohen's theory. Shapiro does discuss Cohen's theory in one of her earlier books. See SHAPIRO, *supra* note 23, at 277 n.69. However, that discussion consists of a short paragraph (in a footnote) and there, as here, Shapiro does not invoke Cohen's theory in her analysis of historical conceptions of probability.

56. See *supra* note 45.

which she and her intellectual compatriots call “mere probability,” “opinion,” and “conjecture.” This conception of species of proof is depicted in Figure 3.

FIGURE 3

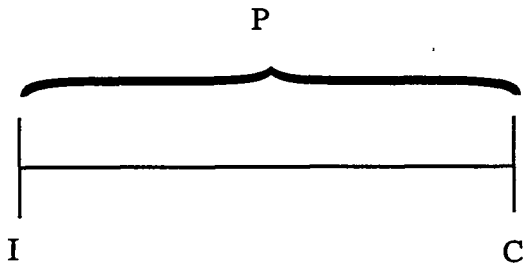


Shapiro's taxonomy of proof presents numerous difficulties. Consider first Shapiro's view of the domain of probability. While almost all modern observers would grant that the "realm of probability is . . . [an] intermediate one," a theorist who believes that the standard probability calculus is the only valid paradigm of rational inference would insist that the outer boundaries of probability are quite different from those shown in Figure 3. She would say that the appropriate distinction is not between (a) conclusions that are merely probably true (or false) and conclusions that are very probably true (or false), but between (b) arguments that make factual hypotheses probable and those that make factual hypotheses either certain or impossible. She would insist that all arguments that fail to produce indubitable conclusions about factual hypotheses fall into the realm of probability. Thus, in her eyes, the realm of probabilistic reasoning is marked off at one end of the scale not by conclusions that are "almost certainly" true, but by conclusions that are indubitably and necessarily true; at the other end of the scale, it is marked off not by weakly supported "suspicion" or "opinion," but by conclusions that are indubitably and necessarily false.

A conventional probabilist envisions species of proof in the manner shown in Figure 4. This figure shows that the only difference among uncertain hypotheses is a difference of degree; some of them are more, or less, probable than some others. In every other respect all uncertain hypotheses — whether labeled "opinion," "probable," or "practically certain" — are exactly the same.

A conventional probabilist would also reject Shapiro's claim that "probability . . . contains no . . . steps, degrees, or levels of probability"

FIGURE 4



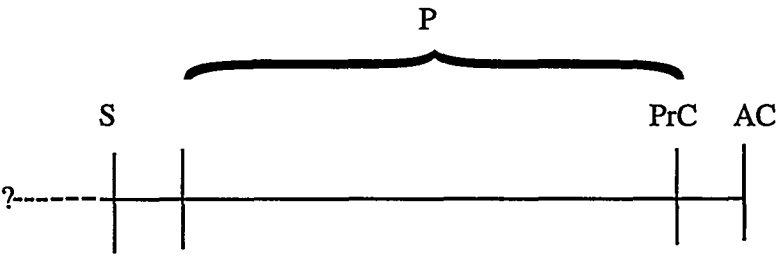
C = "Certainty"

I = "Impossibility"

P = "Probability"

in the interval between "just above opinion" and "just below certainty." As we have already seen, Shapiro places probability somewhere above opinion and somewhere below certainty. Hence, Shapiro conceives of levels of proof in the manner depicted by Figure 5. A

FIGURE 5



AC = "Absolute Certainty"

PrC = "Practical Certainty"

P = "Probability"

S = "Suspicion", "Opinion"

conventional probability theorist would reject the taxonomy of proof shown in Figure 5. He could phrase his objection in one of two ways. On the one hand, he might say that if one type of factual proof can be

distinguished from another type of factual proof, the only possible fundamental distinction between different types of proof is the distinction between probabilistic proof and conclusive proof.⁵⁷ On the other hand, he might say that if there are differences in proof "all along the line," the number of levels of probabilistic proofs is uncountably large.

The gist of the first version of the probabilist's objection is apparent. He claims that no fundamental difference exists between one probabilistic proof and any other. The same type of uncertainty infects every probabilistic proof. The only difference between proofs lies in the different probabilities that they produce.

The second version of the probabilist's objection is essentially the same as the first because, like the first, it rests on the idea that differences in the weights of proofs are expressed as differences in probabilities.⁵⁸ Figure 6 shows how a conventional probability theorist thinks probability behaves. The numbers on the scale in Figure 6 represent different probability values. The possible probability values of hypotheses are the numbers in the interval $\{0,1\}$. This set of numbers is uncountably infinite. Consequently, the claim that in the realm of "mere" probability there are no gradations, degrees, or levels is incorrect; the number of gradations of probability is uncountably infinite.⁵⁹

Practically every student of probability theory or of any other formal theory of uncertainty would have difficulty understanding Shapiro's claim that the probable cause standard, unlike the reasonable doubt standard, fluctuates because probability has no "logical resting

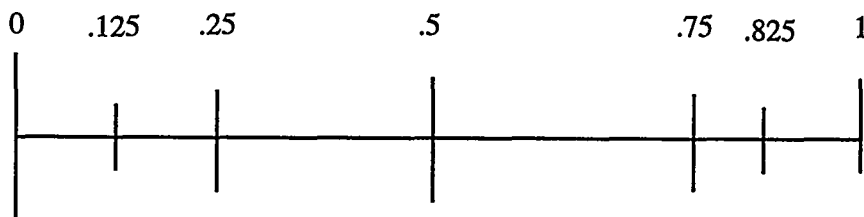
57. In the standard probability calculus conclusive proof is a special case of probabilistic proof. See Jeffrey, *supra* note 8.

58. Inferential arguments can be distinguished on the basis of considerations other than weight. For example, they can be characterized by the type of evidence used or by the structure of the argument. See Peter Tillers, *Webs of Things in the Mind: A New Science of Evidence*, 87 MICH. L. REV. 1225, 1232-1238 (1989) (reviewing DAVID A. SCHUM, EVIDENCE AND INFERENCE FOR THE INTELLIGENCE ANALYST (1987)) (describing Schum's theory of inferential argument).

59. Shapiro's claim that "probability . . . has no agreed upon measure of quantities of probability" is also problematic. Probability theory's "agreed upon measure of quantities of probability" is the set of real numbers ranging from "0" to "1." Moreover, ordinary ways of expressing probabilities — for example, "there is a 60% chance of rain" — often use a similar scale. Shapiro's notion that probability has no measure of probability betrays a failure to appreciate the point of some types of probability talk. In the standard theory of probability, probability is nothing other than a measure or scale of uncertainty. No one should expect that the language of probability itself measures probability. The mathematical language of probability is *used* to measure or express (the magnitude of) uncertainty. Ordinary people often use probabilistic terms in the same way and for the same reason.

Shapiro's claim that "probability contains no *clear* steps [or] degrees," (p. 44, emphasis added) and her claim that the "scale of probabilities ha[s] no *precise* steps," (p. 79, emphasis added) are exaggerations at best. The set of numbers ranging from "0" to "1" can be used to express judgments about uncertainty precisely and clearly. Shapiro's complaint about the alleged absence of a precise measure of probability is odd. Most modern critics of mathematical probability complain that the precision of mathematical expressions tempts people to express probability estimates with too much precision. See, e.g., Tribe, *supra* note 18, at 1331, 1361-65, 1389-91. I see no reason in principle why ordinary ways of expressing probabilities (for example, by the use of percentages) are incapable of being precise.

FIGURE 6



place.”⁶⁰ Any serious student of uncertainty and probability would admit that neither probability nor the logic of probability can possibly determine the “resting place” of probability, much less the resting place of any probabilistically formulated legal requirement such as “probable cause.” Nevertheless, it does not follow that “probable cause” *cannot* have a “resting place,” or settled meaning. Shapiro’s complaint is that the legal standard of probable cause has no clear probability value. However, a probability is not a thing; a probability value reflects or expresses a judgment. Although nothing in probability theory determines the choice of a probability value for probable cause, nothing in the notion of probability *prevents* the assignment of a specific probability value to the legal requirement of probable cause. Shapiro has placed the blame for the law’s fuzziness about the meaning of probable cause at the wrong doorstep.

The last criticism (and criticisms like it) could be made by innumerate students of evidence. No special expertise in mathematical probability is required to see that nothing in the bare idea of the uncertainty of facts prevents courts or legislators from stipulating precisely how much probability (for example, of guilt) is required for a showing of probable cause. Innumerate as well as numerate people can readily understand that the true question is not whether the law *can* be precise. The true question is whether the law *ought* to be precise or *why* the law has not been precise about the meaning of probable cause.⁶¹

I have examined various criticisms of Shapiro’s views of the probable cause and reasonable doubt standards. These criticisms do not prove that Shapiro’s ultimate conclusions are wrong. For example,

60. The notion that probability has no natural resting place is central to Shapiro’s extended discussion of the instability of the probable cause standard through the centuries. Shapiro even attempts to explain the vagaries of modern American probable cause jurisprudence on this basis. See pp. 107-11.

61. Shapiro refers to judicial opinions asserting that probable cause eludes definition. See, e.g., p. 147 (quoting *Illinois v. Galic*, 462 U.S. 213, 231-32 (1983)). However, when courts say that “probable cause” resists definition, they are not necessarily saying that it is logically impossible to attach a probability value to probable cause. Instead, they may be saying that it is imprudent to require a single probability value for all probable cause problems, that it is better for the law to use fuzzy probability values, or that it is better to allow the trier of fact to choose a probability value for probable cause. A footnote in *Beyond Reasonable Doubt* refers to material that suggests these rationales, but Shapiro does not explore them. See p. 273 n.138.

they do not conclusively demonstrate that there is no difference in kind between the reasonable doubt and probable cause standards. However, the purpose of my discussion is not to prove that Shapiro reached the wrong conclusions. My purpose is to show that an intellectual historian who wishes to address problems of evidence in today's world must take into account how today's students of evidence tend to think about probability and proof. For example, although it is possible that the distinction between "practical certainty" and "probable cause" is tenable, a convincing case for the distinction can be made only if the sorts of objections and criticisms described above are confronted directly. My claim is that Shapiro did not do this.

V. INTERROGATING HISTORY

Beyond Reasonable Doubt is an example of painstaking historical scholarship. Nonetheless, there is something wrong with Shapiro's book. Although Shapiro devotes careful attention to details, she sometimes ignores *important* details. The reason is that she sometimes fails to ask the right questions.

A hypothesis is a kind of question; a hypothesis is a question in search of evidence. One of Shapiro's hypotheses is that the conception of evidence and proof now found in the law of evidence is the result of the diffusion of an empiricist epistemology. According to Shapiro, one of the hallmarks of this epistemology is the view that empirical knowledge is a matter of probabilities and degrees of probabilities. Shapiro does an admirable job of gathering and presenting the evidence that supports her hypothesis about the influence of empiricism on the law of evidence. Unfortunately, other hypotheses or questions about empiricism may be far more interesting than the general hypothesis that Shapiro normally addresses.

Most contemporary debates about evidence and proof concern eddies in the stream of probabilistic empiricism rather than the question of the existence or validity of the stream itself. Although Shapiro's book is not about contemporary evidence scholarship, debates and discussions about proof in 1793 as well as in 1993 may have centered more on wrinkles in empiricist epistemology than on the general question of the validity of empiricism.⁶² Shapiro fails to explore some very important details in the historical record. I think this is because her theoretical "frame of discernment"⁶³ is too coarse. I offer three

62. I do not claim that Shapiro ignores all wrinkles in the empiricist tradition. That is plainly not true. For example, in one of her earlier studies Shapiro identified a variety of connotations historically associated with the term "probability." See, e.g., SHAPIRO, *supra* note 23, at 3-6, 37-44, 189.

63. The phrase "frame of discernment" is Glenn Shafer's. See GLENN SHAFER, *A MATHEMATICAL THEORY OF EVIDENCE* 114-40, 172-95 (1976). While Shafer emphasizes how evidence serves to refine coarse hypotheses, my emphasis here is on how the questions and hypotheses we have in mind affect our ability to see interesting evidence.

illustrations.⁶⁴

The first illustration involves the reasonable doubt standard. As discussed above,⁶⁵ Shapiro believes that the reasonable doubt standard requires "practical certainty," which in turn requires proof to the highest probability. I have not argued that the idea of practical uncertainty is necessarily incoherent. However, I have argued that the notion of a highest probability (short of absolute certainty) is incoherent if one assumes that proof is a matter of degrees. Assume I am right. The question now becomes whether the idea of practical certainty can be salvaged not by a frontal assault on the idea that all facts are uncertain to some degree, but by confession and avoidance. In one discussion Shapiro supplies a quotation that provides food for thought about this question.

Shapiro quotes Greenleaf as saying that "satisfactory evidence" of guilt is evidence sufficient "to satisfy the mind and conscience of a common man, and so to convince the jurors, that he would venture to act upon that conviction, in the matters of highest concern and importance to his own interest" (p. 38). In this passage Greenleaf seems to say that the legal requirement of practical or moral certainty demands not the highest possible degree of probability, but the degree of probability that would warrant a wager that the fact at issue is true.⁶⁶ Greenleaf's belief — that jurors should make a decision about the sufficiency of the evidence in a criminal case by asking themselves whether they would be willing to stake important interests of their own on a wager that the accused is guilty — might warm the hearts and stir the minds of students of evidence who have a yen for decision theory.

Decision theorists insist that the problem of choice always involves *two* questions: (a) What are the odds?, and (b) What's at stake? They believe that the second question is as important as the first when a decision must be made and when the options available to the decision maker have consequences. Although Greenleaf plainly wanted to tilt the scales in favor of the accused, Greenleaf's statement suggests that he believed that jurors can locate reasonable doubt and practical cer-

64. I do not claim that the conjectures I describe below are supported by the historical evidence. I only claim that Shapiro's description of historical source material suggests that an examination of some of the conjectures I describe might have been productive.

65. See *supra* notes 47-48 and accompanying text.

66. Cf. DASTON, *supra* note 9.

In the latter half of the seventeenth century, . . . another strain of probabilistic reasoning emerged in the writings of proponents of rational theology and the new natural philosophy. . . . These writers, including Robert Boyle, John Wilkins, Joseph Glanvill, Marin Mersenne, Pierre Gassendi, and Hugo Grotius, simultaneously insisted upon the incorrigible uncertainty of almost all human knowledge and on our ability to nonetheless attain to inferior degrees of "physical" and "moral" certainty. . . . A proof for a hypothesis in natural philosophy or the precepts of Christianity need not achieve mathematical rigor, but only that threshold of certainty sufficient to persuade a reasonable man to act in daily life.

Id. at 56-57.

tainty on the appropriate point in the continuum of probabilities if and only if they consider the gravity or importance of the interests at stake. If this is what Greenleaf meant, it is quite possible that he believed that there is no highest probability of guilt, short of certainty. Furthermore, he may have believed that the appropriate question for the jury is not "How much probability makes the guilt of the accused practically certain?" but, "Given the importance of the interests at stake, what degree of probability should we require before we decide to proclaim the guilt of the accused?" Shapiro ignores this possible "decision theoretic" interpretation of Greenleaf's language.

The quotation from Greenleaf is interesting for another reason. Greenleaf's view that jurors should consider the level of probability that they would demand when their own interests are at stake seems to import a moral principle into the jury's decision making process; that is, Greenleaf may be suggesting (or assuming) that the decision about the amount of probability required for conviction depends on moral considerations. This would not seem odd to some modern theorists. One school of decision theory takes the view that a rational decision maker seeks to maximize expected utility.⁶⁷ Some students of decision theory believe that the maximum expected utility decision rule (or, presumably, any similar decision rule) rests on normative, practical, or psychological considerations rather than solely on logical or epistemological grounds.⁶⁸

My second example of the value of modern thinking in historical investigation of conceptions of evidence and proof involves Wigmore. Shapiro mentions a tripartite classification of evidence in William

67. See, e.g., VON WINTERFELDT & EDWARDS, *supra* note 4, at 18 ("[W]e will state without proof the decision rule to which this book is committed: maximization of expected utility."); Ward Edwards, *Influence Diagrams, Bayesian Imperialism, and the Collins Case: An Appeal to Reason*, 13 CARDOZO L. REV. 1025, 1057 (1991) ("In any decision situation in which the stakes are modest relative to the resources of the decision maker, the optimal decision rule is maximization of utility or expected utility."); see also Isaac Levi, *On Indeterminate Probabilities*, in DECISION, PROBABILITY, AND UTILITY 287, 290 (Peter Gärdenfors & Nils-Eric Sahlin eds., 1988).

Bayesians adopt as their fundamental principle of rational choice the principle that an option is admissible only if it bears maximum expected utility among all the feasible options.

Very few serious writers on the topic of rational choice object to the principle of maximizing expected utility in those cases where [a person's] values and credal state can be represented by a utility function unique up to a positive affine transformation and a unique probability function. The doubts typically registered concern the applicability of this principle.

Levi, *supra*, at 290; see also Paul Slovic, *Choice, in THINKING, AN INVITATION TO COGNITIVE SCIENCE* 89, 90 (Daniel N. Oshershon & Edward E. Smith eds., 1990) (describing the maximum expected utility principle but not endorsing it for all purposes).

68. See, e.g., David T. Wasserman, *Comment on Edwards: Ward Edwards and the New Bayesian Software*, 13 CARDOZO L. REV. 1075, 1077-79 (1991). Decision theorists who say that maximum expected utility is a normative principle as well as a descriptive principle do not necessarily agree with Wasserman. In the context of discussions of decision theory, the term "normative" is ordinarily a synonym for "rational" or "coherent," but not "moral." See, e.g., Slovic, *supra* note 67, at 91, 96 ("normative study" of choice and "model for rational choice" treated as synonymous).

Lambarde's handbook for justices of the peace, which was published in the seventeenth century (pp. 153-54). This tripartite schema arranges evidentiary events on the basis of their temporal relationship to the matter in issue, with events preceding the matter falling into one group, events accompanying the matter into a second, and subsequent events into a third. According to Shapiro, Lambarde borrowed this tripartite classification from the classical rhetorical tradition (pp. 152-53). However, Shapiro says nothing about the possible theoretical significance of Lambarde's tripartite classification.

Wigmore utilized a tripartite schema remarkably similar to Lambarde's.⁶⁹ However, modern evidence scholars, who generally still have great respect for Wigmore, rarely mention his tripartite schema. Apparently they do not think it is important or useful. The similarity between Lambarde's and Wigmore's tripartite classification schema presents the interesting and possibly important question whether Wigmore's classification is merely an unexpected vestige of the rhetorical tradition or whether it is something more. Wigmore's tripartite classification of evidence challenges the unspoken assumption of some modern theorists of evidence that the logic of rational argument about evidence is relational but never causal.⁷⁰ Moreover, Wigmore's schema rests on the important insight that *evidence* may be causally related to "facts in issue."⁷¹ If Shapiro had paid more attention to modern trends in evidence scholarship, she might have realized that the similarity between Lambarde's and Wigmore's tripartite classifications would interest students who believe that causal reasoning plays a role in inference.

My third and final example also involves Wigmore. Shapiro quite appropriately emphasizes the importance of the empiricist conception of probability in the Anglo-American legal consciousness after the late seventeenth century. However, although the idea that proof concerns probabilities and degrees is undeniably important, the mere idea that all facts (or almost all facts) are uncertain leaves unanswered many important questions about the nature of factual proof. Shapiro argues that Wigmore, like other English and American treatise writers, believed that proof involves probabilities (p. 40). This argument is correct. But it does not go far enough.

Wigmore did not embrace the calculus of chances. Indeed, as far I can determine, he never even referred to it. This is not particularly

69. 1A WIGMORE, *supra* note 15, § 43 (describing "prospectant," "concomitant," and "retrospectant" evidence). This tripartite classification is the basis for the organization of Wigmore's entire treatise on the law of evidence. See the table of contents in any volume of Wigmore's treatise.

70. See Peter Tillers & David A. Schum, *Charting New Territory in Judicial Proof*, 9 CARDOZO L. REV. 907, 942-49 (1988) (explaining the distinction between relational analysis and nonrelational evidence marshalling methods).

71. See Tillers, *supra* note 58, at 1248-49.

surprising or important. What is interesting is that Wigmore almost certainly did not believe that the *logic* of the calculus of chances captures the essence of rational inference and proof. Wigmore's intellectual ancestors were not mathematical probability theorists such as James Bernoulli, who discovered the "law of large numbers," or the Reverend Thomas Bayes, who formulated the theorem now known as "Bayes' Theorem." Instead, judging by the citations and quotations in Wigmore's justly celebrated treatise on the law of evidence, Wigmore was influenced by figures such as Jeremy Bentham and John Stuart Mill.⁷² Of course, both Bentham and Mill believed that proof is a matter of probabilities. Nevertheless, this does not mean that they thought about uncertainty and probability the same way that British theorists such as Bayes did.

Consider the experimental methods that Mill favored, now known as "Mill's methods." Wigmore thought that Mill's analysis about the nature of empirical proof and scientific investigation was right on target.⁷³ However, unlike Hume (who was surely an "empiricist") and some modern probability theorists, Mill apparently believed that causes exist in the world of nature; Mill assumed that a scientist attempts to uncover causes that actually operate in the natural world.⁷⁴ Wigmore's endorsement of Mill's theory of empirical knowledge distinguishes Wigmore's own epistemological position from other epistemological positions, including "empiricist" positions, in a way that no amount of talk solely about degrees of probability can. Moreover, Mill's methods put a considerable premium on the exhaustiveness or completeness of evidence about a hypothesis in a way that Bayesianism, for example, does not.⁷⁵ Finally, it is possible that Mill's methods, unlike Bayesian logic, interpret grades of proof ordinally rather than cardinally; Mill, like some modern-day Baconians, may have believed that a scientific hypothesis has been "proven" if a sufficiently complete collection of evidence supports that hypothesis without offering any support for an alternative hypothesis.⁷⁶ One of Wigmore's

72. See, e.g., 1A WIGMORE, *supra* note 15, § 33. It is interesting that Shapiro acknowledges the influence of Mill on Sir Fitzjames Stephen. P. 273 n.134. However, Shapiro does not explore the significance of this point.

73. 1A WIGMORE, *supra* note 15, § 24 (quoting from the work of Alfred Sidgwick, a follower of Mill, with approval); see also, e.g., *id.* §§ 30, 31, at 991-92; § 33, at 996-98. Although Wigmore believed that Mill's methods are valid, he did not believe that they are the only valid methods for making judgments about uncertain facts. See, e.g., JOHN H. WIGMORE, *THE SCIENCE OF JUDICIAL PROOF* §§ 340-352 (3d ed. 1937) (Wigmore's description of his "chart method").

74. See, for example, Mill's discussion of plurality of causation in 1 JOHN STUART MILL, *A SYSTEM OF LOGIC, RATIOCINATIVE AND INDUCTIVE* bk. 3, ch. x (Parker, Son & Bourn 5th ed. 1862).

75. See COHEN, *supra* note 44, § 46 (Cohen, who emphasizes the importance of the completeness of relevant evidence, interprets Mill's methods as an example of his own "method of relevant variables.").

76. This is roughly the position taken by Cohen. See Cohen, *supra* note 53.

formulations of the relevance principle suggests that Wigmore had an ordinal conception of the nature of proof.⁷⁷

Shapiro's treatment of Wigmore is a particularly apt example for my purposes. If any evidence scholar from the past still influences evidence scholarship and the law of evidence today, it is Wigmore. The question of Wigmore's intellectual pedigree, therefore, has a direct bearing on contemporary scholarly and judicial discussions and debates about evidence, inference, probability, and proof. If Shapiro's "long view"⁷⁸ of intellectual history is responsible for her failure to wrestle with the details of the ideas of figures such as Wigmore, her view is overly long.

CONCLUSION

Beyond Reasonable Doubt has little to say to modern students of evidence. Shapiro's theoretical brush is too coarse; she fails to make important distinctions among different conceptions of evidence, inference, probability, and proof. The probable cause of the theoretical weaknesses of *Beyond Reasonable Doubt* is Shapiro's failure to pay careful attention to modern discussions of evidence and proof. Although historians are sometimes faulted by their peers for thinking anachronistically about the past, good intellectual history may require immersion in the ideas of the present as well as of the past. Be that as it may, Shapiro's study would have been more interesting to today's students of evidence if she had made a greater effort to structure her historical investigation in terms of present-day theoretical perspectives and concerns. Doing so would not have prevented her from using history to shake modern scholars free from their unconsidered habits of thought and their theoretical prejudices. The peculiar value of historical inquiry is that the mere framing of a question does not predetermine the answer; like any other body of evidence, historical source material always contains surprises.

77. 1 WIGMORE, *supra* note 15, § 32 at 996.

Thus, throughout the whole realm of evidence, . . . the theory of the inductive argument . . . is that the evidentiary fact will be considered *when and only when the desired conclusion based upon it is a more probable or natural, or at least a probable or natural, hypothesis and when the other hypotheses or explanations of the fact . . . are either less probable or natural or at least not exceedingly more probable or natural.*

Id. (emphasis added); cf. 1A WIGMORE, *supra* note 15, § 30, at 986-89 (reviser argues that Wigmore's theory of inference and proof resembles Mill's theory more than it does modern "personalist" theories of inference and probability).

78. See *supra* note 26.